

文章编号: 2095-4980(2014)06-0912-06

基于数据仓库及 OLAP 的鞋类销售系统设计与实现

丁 洁

(陕西工业职业技术学院 信息工程学院, 陕西 咸阳 712000)

摘 要: 传统鞋类销售系统数据收集不全面, 准确性不高, 因而导致决策不科学, 并产生资源浪费等问题。本文通过对鞋类销售系统数据仓库的设计, 利用联机分析处理(OLAP)技术, 探讨了基于数据仓库和联机分析处理技术的鞋类销售分析设计、数据迁移、销售数据分析等问题。实现了鞋类销售系统中的数据多维分析, 具有一定的应用价值。

关键词: 数据仓库; 鞋类销售系统; 联机分析处理; 多维数据库

中图分类号: TN919.6⁺5; TP311.133.1 **文献标识码:** A **doi:** 10.11805/TKYDA201406.0912

Design and implementation for footwear sale system based on data warehouse and OLAP

DING Jie

(Department of Information Engineering, Shaanxi Polytechnic Institute, Xianyang Shaanxi 712000, China)

Abstract: There exist many problems in the traditional footwear sales system, such as its data collection is not comprehensive, and the accuracy is not high, which will lead the unscientific decision-making and therefore cause the resources waste. In this paper, the data warehouse shoe sale system is designed. The footwear sales analysis design, data migration, sales data analysis based on the data warehouse and On-Line Analytical Processing(OLAP) technology are discussed. The multidimensional analysis for footwear sales data in the system is realized.

Key words: data warehouse; footwear sales system; On-Line Analytical Processing; multidimensional database(MDDB)

目前, 多数鞋类厂商的销售系统对于底层数据的收集不全面, 导致了在制定生产、销售计划时缺乏可行和有效的数据依据, 而不能进行科学的决策。如果没有准确的市场分析, 对市场进行精准的了解与定位, 只是盲目扩大生产规模或销售网络, 最终会导致企业资源的浪费。因此, 建立数据仓库架构下, 以信息技术、计算机技术、网络技术为基础的支持联机分析处理(On-Line Analytical Processing, OLAP)的鞋类销售管理信息系统, 实现销售决策分析的科学化、销售管理的规范化十分必要。本文以鞋类销售数据仓库设计为基础, 探讨了基于数据仓库和联机分析处理技术的鞋类销售分析设计、数据迁移、销售数据分析等问题, 实现了鞋类销售系统多维数据的分析与实现。

1 基于数据仓库的销售系统体系结构设计

基于销售数据、当前和历史的详尽营销信息等, 创建数据仓库, 建立元数据, 对信息分析的需求重新部署, 在有针对性地定义和组织数据仓库中的数据基础上, 确保数据的一致性与有效性^[1]。针对鞋类销售系统业务处理过程中诸如: 鞋品管理、订单管理、鞋品流转销售管理、资金结算等方面的数据, 进行清理、转换、抽取、刷新等操作后进入数据仓库, 形成关联性较强的多维数据, 实现对鞋类销售系统中数据的 OLAP, 其分析结果将提供给企业高级营销人员的决策层。结合该系统实际, 本文设计的鞋类管理系统技术体系结构, 如图 1 所示。

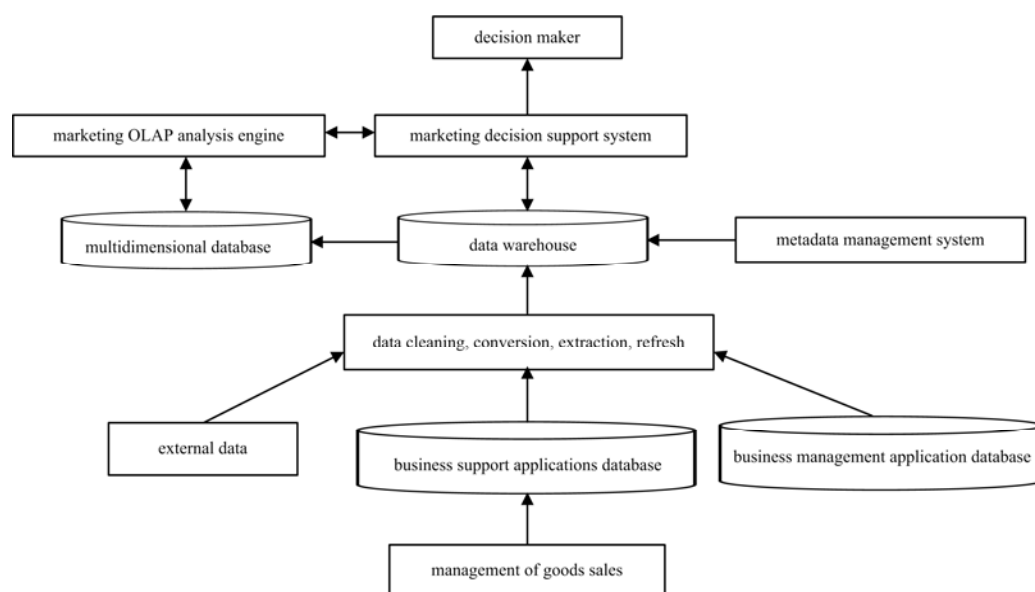


Fig.1 Architecture of footwear management system

图 1 鞋类管理系统体系结构

2 销售系统数据仓库设计

2.1 鞋类销售数据仓库的主题设计

根据鞋类销售系统的实际需求，围绕鞋品、时间、客户和职员等主要信息的相关内涵数据的主题，在该系统数据仓库中创建 4 个维，维表详见表 1。

表 1 数据仓库的维表信息

Table1 Dimension table information of data warehouse

dimension	dimension description	items
footwear	product ID, name, origin, specifications, packaging, color, texture	product ID
time	time ID, date, year, month, day, week, quarter	time ID
customer	the customer number, contact name, address, telephone, occupation	customer ID
employee	employee number, name, sex, age, culture degree, nationality, address, telephone, departments, wage	employee ID

2.2 概念模型设计

4 个维表的建立，实际上就是对原有系统的数据进行了归纳分析，搞清楚原业务系统中有什么数据，关系是如何组织的，怎样进行分布等等。在此基础上，对原有销售数据进行重新组合，形成新的数据集合(数据仓库)。考虑到概念模型设计不受具体技术条件限制，本系统建立的数据仓库的关系型数据库结构采用星型模型，由 1 个事实表和 4 个维度表组成，维度表提供决策支持，如图 2 所示。

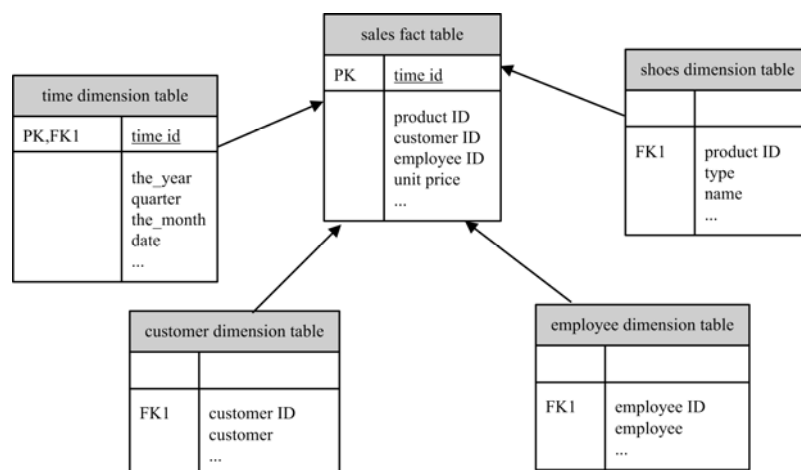


Fig.2 Footwear star model

图 2 鞋类销售星型模型

2.3 鞋类销售数据仓库逻辑建模

1) 确定事实表和维度表

鞋类商品的销售环节部分是销售决策系统中最应予以关注的，按照上述设计的鞋类销售系统的星型模型架构，建立数据仓库中的事实表和维度表十分关键。因此，将以销售信息为主题的部分作为事实表，相关的信息如：鞋类、时间、客户、职员等的信息作为维度表。

2) 设计事实表

事实表的设计要尽可能围绕销售信息这一主题，应包含有商业事务的详细信息。所以，围绕鞋类、时

间、客户和职员的信息设计的事实表字段及数据类型如表 2 所示。因考虑系统实用性和测试的因素,为了减少查询和处理的时间,通过降低列的数量、列的大小(如将 char(60)统一为 Varchar(60))来降低表的容量。

表 2 销售事实表

Table2 Sales fact table		
items name	description	data type
time code	time ID	int
product code	product ID	int
customer code	customer ID	int
employee code	employee ID	int
footwear prices	unit price	numeric8,2
the purchase price	purchase price	numeric8,2
number of sale	number	numeric10,2
value of sales	value of sales	numeric10,2
cost of sales	cost of sales	numeric10,2
sales profit	profit	numeric10,2

3) 设计维度表

设计维度表的主要目标是将非正规化的参考事实表数据填入一个单独的表。通过直接参考事实表,而不是通过其他维度表间接参考维度表,最小化表连接的数量,提高系统的性能。维度表如表 3~表 6 所示。

表 3 时间维度

Table3 Time dimension		
items name	description	data type
time code	time ID	int
gregorian calendar date	the-date	date
the English week	the-day	character 15
Chinese month	the-month	character 4
Chinese year	the-year	character 6
day of month	day-of-month	int
week of year	week-of-year	int
month of year	month-of-year	int
Chinese quarter	quarter	character 6

表 4 客户维度

Table4 Customer dimension		
items name	description	data type
customer code	customer ID	int
customer name	customer	varchar 40
customer type	type	varchar 10
customer region	region	varchar 10

表 5 鞋品维度

Table5 Shoes dimensions		
items name	description	data type
product code	shoe ID	int
product name	shoe name	varchar 30
product specifications	shoe size	varchar 6
product type	shoe type	varchar 6
product origin	shoe origin	varchar 30
product packaging	shoe packaging	numeric 4,0
product color	shoe color	varchar 10
product texture	shoe texture	varchar 50
product unit	shoe unit	varchar 2

表 6 职工维度

Table6 Employees dimension		
items name	description	data type
employee code	employee ID	int
employee name	name	varchar 8
employee sex	sex	varchar 2
employee age	age	varchar 10
employee nationality	nationality	varchar 2
employee native place	native place	varchar 12
the level of education	the level of education	varchar 6
date of birth	date of birth	date
ID number	ID number	varchar 20
telephone	telephone	varchar 14
department	department	varchar 20

4) 实现数据仓库设计

在事实表和维度表设计并建立的基础上,利用 SQL Server 2005 的企业管理器,进行数据仓库的创建工作。在具体创建过程中,本文采取了在每一个维度表的关键字列上建立一个索引,用来加速数据仓库的检索速度,减少数据检索时间,有效提高系统效率。

3 基于数据仓库的销售系统 OLAP 实现

3.1 鞋类销售系统实现 OLAP 分析的架构设计

经过上述数据仓库设计与实现,就可在此基础上完成从原处理系统中的数据按照主题进行抽取、转换,最终能够加载到鞋类销售的数据仓库中^[2-4]。通过这一途径,决策层可以采取企业报表或联机多维分析等两种方式来访问数据仓库,从中调用数据信息,全面、统一地实现使用者对销售系统信息的掌控,在数据仓库服务下,确保了获取数据的唯一性和准确性^[5],如图 3 所示。

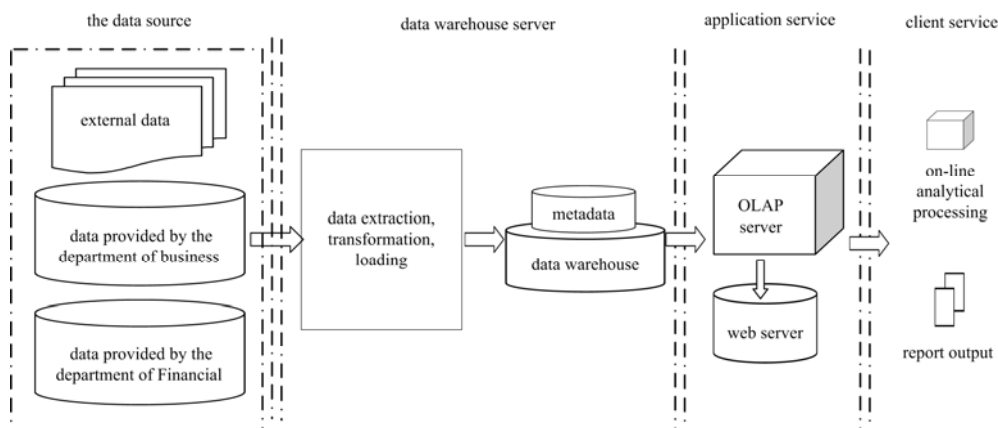


Fig.3 Footwear sales chart
图 3 鞋类销售架构图

3.2 鞋类销售系统多维数据库(MDDB)设计

在数据仓库上实现 OLAP 之前,首先要创建多维数据库^[6]。因此,在上述维度表设计的基础上,确定 MDDB 的维数和维的内涵是构建 MDDB,搭建数据仓库 OLAP 分析数据引擎的关键。结合本文决策层的需求和实际,对鞋类、客户、时间、职工等 4 个维进行了细化,本系统多维数据集的维结构如图 4 所示。同时,本文根据实际销售的需求,灵活定义各维的汇总层次,并在形成的多层次维情况下,按照维的层次划分并形成每一维的汇总路径。依据形成的路径,当数据沿其中任何一条路径自下而上地分析时,就实现了钻取分析(Drill down);如果将某一路径与不同方向的其他路径作任意组合,则是切片与切块分析(Slice & Dice)。

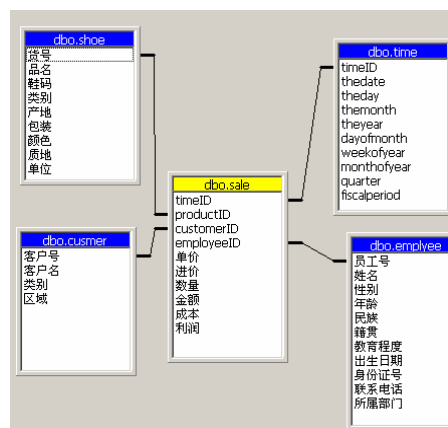


Fig.4 Footwear sales cube-dimensional structure
图 4 鞋类销售多维数据集的维结构

3.3 基于数据仓库的鞋类销售系统的 OLAP 实现

1) 系统的 OLAP 实现

在创建了数据仓库、MDDB 等操作后,即可通过联机分析处理来分析数据,为决策层提供销售依据^[7]。通常可以利用引导分析的手段,即根据相应管理人员所给出的条件,得到企业高层需要的信息,以满足决策层的认知水平和便于识别的表象;或者是采取数据挖掘的手段,从大量数据集合中,有效地发现对商品销售有利的实时信息;而多维分析的手段,则是对以多维形式组织起来的数据,进行钻取、切片与切块等各种分析,以便对数据进行归纳和解析,使企业管理人员或决策层能够从多个角度和不同的侧面观察数据仓库中的数据,最终了解数据包含的信息及内涵^[8]。本文以鞋类销售系统中的销量为例,实现了联机分析多维分析与实现,见图 5~图 7。



Fig.5 Processing cube
图 5 多维数据集的处理

+类别	MeasuresLevel	单价	进价	数量	金额
- 所有客户		791.00	560.00	3.00	720.00
+ 高级		350.00	260.00	1.00	320.00
+ 普通		291.00	220.00	1.00	280.00
+ 中低		150.00	80.00	1.00	120.00

Fig.6 Review of footwear sales data cube
图 6 鞋类销售多维数据集集中数据浏览

+类别	MeasuresLevel	单价	进价	数量	金额
- 所有客户		291.00	220.00	1.00	280.00
+ 高级					
+ 普通		291.00	220.00	1.00	280.00
+ 中低					

Fig.7 Sales analysis for 2013-5-9 of footwear cube
图 7 鞋类销售 2013-5-9 多维数据集销量分析

SQL Sever 2005 作为开发阶段时常采用的工具,采取数据仓库挖掘技术的同时对某鞋类企业厂商的各个数据进行查询,把分散的数据按照相应的规则组织起来,提供了可靠的信息支撑。以 WEB 为基础的分发方式、数据展现方式,从某种意义上来说有助于外部人员随时访问系统,使管理层有效地利用科学的管理工具实现对未来的预测和分析。这样可以实现决策科学化、数据客观化、管理民主化,同时实现管理上的技术创新。

2) 系统测试

通过设计改造数据仓库鞋类销售系统,并实现在此基础上的联机分析处理,实现了传统的基本数据单维结构向多维结构的转变。由以往的数据查询单调功能、部门单维度查询功能,转变为以销售立方体为基础的各个关系数据库,通过 OLAP 技术,实现销售业务平台,具有综合平台多任务、多维度的查询功能、分析功能、数据钻取功能、图表出具功能,减少了系统对于数据口径不一致导致的数据分析失误差率问题。具体完成了相关测试,测试结果见表 7。

表 7 基于数据仓库的某鞋类销售系统测试报告

Table7 Test report for shoe sales system based on data warehouse

number	test content	test results
1	configuration testing	reach the standard
2	time response test	When the database query time is more than two months, the query speeds low down.
3	response test	The system runs on a CPU 2.6G, memory 8G, hard disk 2TB, whose capacity can meet the requirements.
4	parallel test	Run 200 users at the same time, and the system runs smoothly.
5	the graphic expression test	It can achieve a two-dimensional table data completely.
6	backup test	The system sets up automatic save function, which can report errors when encountering crash circumstances.
7	security test	basically meet the needs of different levels of management authority

4 结论

基于联机分析处理技术的操作业务系统的开发和实现,是零售业市场上涉及不深或者是涉及较少的领域,本文通过建立鞋类商品销售数据仓库,利用联机分析的快速数据统计为鞋类企业厂商提供决策支持,引入 OLAP 技术来辅助鞋类销售系统数据分析的思路,提出了基于数据仓库技术构架下联机处理分析系统的一种解决方案,也给出了具体数据仓库的建立模型及分析数据的方式和步骤,测试结果良好。为企业方便快捷地找到销售异常的原因,预测某款鞋类销售的未来走向以及提高管理工作效率和质量提供了依据和支撑。

参考文献:

- [1] 徐顺飞,吴晓红,何小海,等. 岩心库综合管理系统的多数据库兼容设计[J]. 太赫兹科学与电子信息学报, 2012,10(6): 730-734. (XU Shun-fei, WU Xiao-hong, HE Xiao-hai, et al. Compatibility design of multi-database for core library integrated management system[J]. Journal of Terahertz Science and Electronic Information Technology, 2012,10(6):730-734.)
- [2] 秦秀洁. 数据挖掘流程改进研究[J]. 河南科学, 2013(6):868-872. (QIN Xiu-jie. The data mining process improvement research[J]. Henan Science, 2013(6):868-872.)
- [3] 王志宏,顾新建. 基于数据仓库的产品族供应链管理系统[J]. 浙江大学学报, 2009,43(7):30-32. (WANG Zhi-hong, GU Xin-jian. Product family data warehouse system based on supply chain management[J]. Journal of Zhejiang University, 2009,43(7):30-32.)
- [4] 张春阳. 基于数据仓库的决策支持系统的构建[J]. 计算机工程, 2002,28(4):249-251. (ZHANG Chun-yang. The construction of data warehouse based decision support system[J]. Computer Engineering, 2002,28(4):249-251.)

(下转第 926 页)